

AD-A194 003

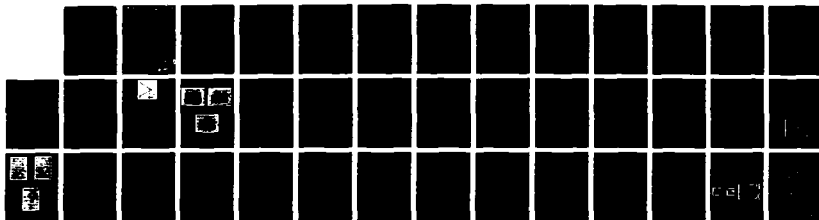
ACOUSTOOPTIC PROCESSING OF TWO DIMENSIONAL SIGNALS
USING TEMPORAL AND SPA. (U) CALIFORNIA INST OF TECH
PASADENA DEPT OF ELECTRICAL ENGINEER.

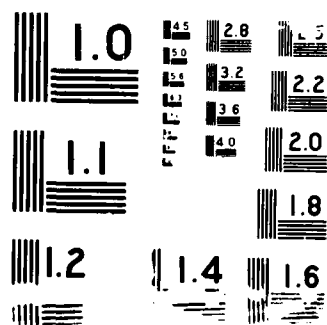
1/1

UNCLASSIFIED

D PSALTIS ET AL. 28 FEB 88 AFOSR-TR-88-0410 F/G 17/9

NL





~~UNCLASSIFIED~~

SEE INSTRUCTIONS ON REVERSE PAGE

DOCUMENTATION PAGE

DTIC FILE COPY

Form Approved
OMB No. 0704-0188

AD-A194 803

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE		1b. RESTRICTIVE MARKINGS	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release, distribution unlimited	
6a. NAME OF PERFORMING ORGANIZATION California Inst. of Tech		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 88-0410	
6b. OFFICE SYMBOL (if applicable)		7a. NAME OF MONITORING ORGANIZATION AFOSR/NE	
6c. ADDRESS (City, State, and ZIP Code) Dept of Electrical Eng California Inst of Tech Pasadena, CA 91125		7b. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling AFB, DC 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-85-0332	
8b. OFFICE SYMBOL (if applicable) NE		10. SOURCE OF FUNDING NUMBERS	
8c. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling AFB, DC 20332-6448		PROGRAM ELEMENT NO. 61102F	TASK NO. 2305
		WORK UNIT ACCESSION NO. B1	
11. TITLE (Include Security Classification) ACOUSTOOPTIC PROCESSING OF TWO DIMENSIONAL SIGNALS USING TEMPORAL & SPATIAL INTEGRATION			
12. PERSONAL AUTHOR(S) Demetri Psaltis, Dave Brady, Scott Hudson, Fai Mok, Jeff Yu			
13a. TYPE OF REPORT Annual	13b. TIME COVERED FROM 1 Mar 87 to 28 Feb 88	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT 36
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The goal of this project is the development of optical signal processing architectures and techniques that are suitable for processing information in two dimensions. We have applied our methods to synthetic aperture radar (SAR), image recognition, two dimensional spectrum analysis of one dimensional signals and adaptive phased arrays. In previous years we have developed several specific acoustooptic architecture in each category. In this report we describe our recent results in the areas of radar imaging, image correlation, and integrated optical signal processing			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Lee Giles		22b. TELEPHONE (Include Area Code) (202) 767-4933	22c. OFFICE SYMBOL NE

DD Form 1473, JUN 86

Previous editions are obsolete.

UNCLASSIFIED

88 3 02 116

UNCLASSIFIED

DTIC
ELECTE
MAY 04 1988
S
Q
E
D

Annual Report

**ACOUSTOOPTIC PROCESSING OF
TWO DIMENSIONAL SIGNALS USING
TEMPORAL AND SPATIAL INTEGRATION**

Demetri Psaltis, Dave Brady
Scott Hudson, Fai Mok, Jeff Yu

Grant AFOSR-85-0332

Submitted to:
Dr. Lee Giles
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC

Principal Investigator

Demetri Psaltis
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA, 91125

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



TABLE OF CONTENTS

I INTRODUCTION	1
II INVERSE SYNTHETIC APERTURE RADAR : IMAGING METHOD AND OPTICAL IMPLEMENTATION	3
II.1 Introduction	3
II.2 Fundamentals of Radar Imaging	3
II.3 Motion Compensation	4
II.4 ISAR Imaging in Two Dimensions	5
II.5 Effects of Neglecting Aspect Compensation	7
II.6 Learning Aspect Compensation	8
II.7 Optical Implementation	10
III CAPACITY OF OPTICAL CORRELATORS	14
III.1 Introduction	14
III.2 Capacity of Linear Filters	14
III.3 Capacity of Shift Invariant Filters	15
III.4 Capacity of Binary Filters	17
III.5 Capacity of the Volume VanderLugt Correlator	17
III.6 Conclusion	19
IV PHOTOREFRACTIVE INTEGRATED OPTICAL VECTOR MATRIX MULTIPLIER	26
IV.1 Introduction	26
IV.2 Matrix Implementation Via Integrated Volume Holograms	26
IV.3 Recording the Matrix	29
IV.4 Experimental Results and Discussion	31
IV.5 Conclusion	32

I

INTRODUCTION

The goal of this project is the development of optical signal processing architectures and techniques that are suitable for processing information in two dimensions. We have applied our methods to synthetic aperture radar (SAR), image recognition, two dimensional spectrum analysis of one dimensional signals and adaptive phased arrays. In previous years we have developed several specific acoustooptic architectures for each application area and we have experimentally demonstrated at least one architecture in each category. In this report we describe our recent results in the areas of radar imaging, image correlation, and integrated optical signal processing.

In the field of radar imaging we have begun to investigate methods for performing inverse synthetic aperture radar (ISAR) imaging. In normal synthetic aperture radar, radar platform motion is used to create a long synthetic aperture that allows high azimuth resolution to be achieved. If there is any target motion during the synthesizing of the aperture the resulting image will be degraded. Consequently SAR imaging is only applicable to stationary targets. ISAR, however, exploits target motion to create a synthetic aperture without the need for radar platform motion. Unfortunately the data recorded by the synthetic aperture is determined not only by the nature of the target but also by the target's motion. The effects of the target's motion on the recorded data must be compensated for if a faithful image of the target is to be produced. Currently ISAR imaging techniques make restrictive assumptions about the target's motion. We have developed an iterative approach which is applicable to more general kinds of target motion than current techniques. We discuss our approach to ISAR imaging in section II.

VanderLugt correlators have been used for a long time for optical pattern recognition. Recently there has been a great deal of interest in using VanderLugt correlators in pattern classification systems. These systems determine whether an input image is a member of one of two classes, with each class being composed of many images. Typically a reference filter is formed from a linear combination of these images. When an image is input to this filter the presence or absence of a correlation peak determines which of the two classes the image belongs to. The capacity of a correlator is the maximum number of images that can be stored in the reference filter without the system misclassifying an image. In section III we examine the capacity of VanderLugt correlators using both planar and volume optical recording media. In addition we examine the capacity when the filter is binarized.

Vector-matrix multipliers are fundamental components of many signal processing systems such as neural networks and any system which must perform general linear transformations. Implementing vector-matrix multipliers with bulk optics has proven quite successful, especially where computational speed and low power consumption are of prime importance. Despite this success integrated optical implementations offer certain advantages over bulk systems such as smaller size, lighter weight, and ease of mass fabrication.

Additionally integrated optical vector-matrix multipliers are planar devices which allow use of the third dimension for optical programming of the matrix. In section IV we present an architecture for an integrated optical vector matrix multiplier using a photorefractive medium for the matrix. We also discuss techniques for writing the matrix using unguided light from out of the plane of the device.

II

INVERSE SYNTHETIC APERUTURE RADAR : IMAGING METHOD AND OPTICAL IMPLEMENTATION

II.1 Introduction

Radar has proven over many years as an effective means of detecting and tracking targets. A very desirable additional feature is the ability to identify the target being tracked. The most straight-forward approach to target identification is to produce an image of the target which can then be presented to a pattern recognition system. Unfortunately radars, due to practical constraints on their physical apertures, are generally not able to obtain enough information from a single radar echo to produce an image. Instead they must *synthesize an aperture* by recording over time, as either the target or the radar moves, a series of echos from the target until enough information is obtained to produce an image. The recorded echos contain information not only about the target's reflectivity, which is what we wish to compute in the imaging process, but also the about the relative motion between radar and target. The effects of the latter can interfere with the imaging process and must be compensated for if a well-focused image of the target is to be produced.

II.2 Fundamentals of Radar Imaging

When viewed from a small range of angles a target can be characterized by a *reflectivity function* $f(\vec{r})$ such that if a field $E(t)$ is incident from a radar onto a volume d^3r , centered at a point \vec{r} on the target, then a field $f(\vec{r})E(t)d^3r$ is reflected back towards the radar (we use a coordinate system centered and fixed on the target). Imaging a target is the process of computing $f(\vec{r})$.

If the radar sits at the point $\vec{\rho}$ and transmits a field

$$e^{j2\pi\nu_0 t}$$

then the reflected field received at the radar is given by

$$e^{j2\pi\nu_0 t} E(\nu_0) = e^{j2\pi\nu_0 t} \int f(\vec{r}) e^{-j2\pi\nu_0 2|\vec{\rho}-\vec{r}|/c} d^3r \quad (II.1)$$

which is the sum of the reflected fields from all points on the target, each of which has a phase shift due to the $2|\vec{\rho}-\vec{r}|/c$ time it takes the signal to go from the radar to the point \vec{r} and back. If $\rho \equiv |\vec{\rho}|$, the *gross-range*, is large enough to put the radar in the target's far-field then the far-field approximation

$$|\vec{\rho}-\vec{r}| \approx \rho - \vec{e}_\rho \cdot \vec{r} \quad (II.2)$$

is valid, where \vec{e}_ρ – *the aspect* – is the unit vector in the radar's direction. Using (II.2) in (II.1) and mixing out the temporal carrier $\exp(j2\pi\nu_0 t)$, we have

$$E(\nu_0) = e^{-j2\pi 2\nu_0 \rho/c} \int f(\vec{r}) e^{j2\pi [2\nu_0 \vec{e}_\rho/c] \cdot \vec{r}} d^3 r. \quad (II.3)$$

Using the definition of the Fourier transform

$$\hat{f}(\vec{u}) = \int f(\vec{r}) e^{j2\pi \vec{u} \cdot \vec{r}} d^3 r$$

(II.3) can be written as

$$E(\nu_0) = e^{-j2\pi 2\nu_0 \rho/c} \hat{f}(2\nu_0 \vec{e}_\rho/c).$$

Thus, apart from a gross-range-induced phase distortion, the radar has obtained a sample of the target's spatial spectrum. The location of the sample in the target's Fourier space is determined by the radar's frequency and aspect. If instead of a pure sinusoid the radar transmits a signal of bandwidth β around a center frequency ν_0 then the target's spectrum will be sampled along that portion of the ray \vec{e}_ρ from $2(\nu_0 - \beta/2)/c$ to $2(\nu_0 + \beta/2)/c$ in magnitude. If additionally some relative motion between the target and radar causes ρ and \vec{e}_ρ to become functions of time – $\rho(t)$ the *gross-range history* and $\vec{e}_\rho(t)$ the *aspect history* respectively – then

$$E(\nu_0 + \nu, t) = e^{-j2\pi 2(\nu_0 + \nu)\rho(t)/c} \hat{f}(2(\nu_0 + \nu)\vec{e}_\rho(t)/c) \quad (II.4)$$

for $-\beta/2 \leq \nu \leq \beta/2$. Hence by exploiting wide-band signals and radar/target motion a region of the target's Fourier space can be recorded. This process is referred to as *synthesizing an aperture*. Computing $f(\vec{r})$ is then a matter of putting this data in a suitable form and inverse Fourier transforming.

II.3 Motion Compensation

If we know $\hat{f}(\vec{u})$, the Fourier transform of $f(\vec{r})$, in a region of Fourier space we can inverse transform and obtain $f(\vec{r})$ to within the resolution determined by the finite extent of the region. Unfortunately the data $E(\nu_0 + \nu, t)$ is not recorded as a function of the target's spatial frequency, \vec{u} , but as a function of radar frequency and time, $\nu_0 + \nu$ and t respectively. To figure out the value of $\hat{f}(\vec{u})$ at some position \vec{u}_1 in Fourier space from the recorded data we need to find ν_1 and t_1 such that

$$2(\nu_0 + \nu_1)\vec{e}_\rho(t_1)/c = \vec{u}_1. \quad (II.5)$$

so that

$$E(\nu_0 + \nu_1, t_1) = e^{-j2\pi 2(\nu_0 + \nu_1)\rho(t_1)/c} \hat{f}(\vec{u}_1).$$

In addition we need to know the value of $\rho(t_1)$ so that we can eliminate the phase distortion and arrive at

$$\hat{f}(\vec{u}_1) = e^{j2\pi 2(\nu_0 + \nu_1)\rho(t_1)/c} E(\nu_0 + \nu_1, t_1).$$

The frequency ν_1 can be solved for easily from (II.5)

$$\|\vec{u}_1\| = 2(\nu_0 + \nu_1)/c \Rightarrow \nu_1 = \frac{1}{2}c\|\vec{u}_1\| - \nu_0.$$

To find t_1 from (II.5) we need to solve

$$\vec{e}_\rho(t_1) = \vec{u}_1/\|\vec{u}_1\|$$

which requires knowledge of the aspect history $\vec{e}_\rho(t)$ just as removal of the phase distortion requires knowledge of the gross-range history $\rho(t)$. Removing gross-range-induced phase distortion and mapping from radar frequency and time to position in the target's Fourier space are referred to as *motion compensation*. We will refer to the former as *gross-range compensation* and to the latter as *aspect compensation*.

The term *synthetic aperture radar* (SAR) is used to describe situations in which the target is stationary and relative motion between radar and target is due to radar motion alone. Since the radar's motion is generally controllable and/or measurable both the gross-range and aspect histories are known and hence motion compensation can be performed and an image of the target produced.

Inverse synthetic aperture radar (ISAR) describes situations in which the target is moving so relative motion is due, at least in part, to target motion. This is the kind of imaging we will discuss in this paper. Since target motion is generally *not* controllable and *not* measurable we lack the complete knowledge of the gross-range and aspect histories needed to perform motion compensation. Thus ISAR imaging is not possible unless we can measure $\rho(t)$ and $\vec{e}_\rho(t)$ or figure out some way to compute them during the imaging process. Because radars are range measuring devices it is possible to measure $\rho(t)$ and techniques have been developed to perform gross-range compensation [1,2]. Therefore in what follows we will assume this has been done and concentrate on the problem of performing aspect compensation.

II.4 ISAR Imaging in Two Dimensions

We are going to examine the special case in which target motion is such that $\vec{e}_\rho(t)$ is confined to a single plane through the origin. This accurately describes a great many cases of practical interest – e.g. targets moving in straight lines, targets rotating about some fixed axis – and is the case commonly treated in the literature [1,2,3]. By the choice of coordinate axes, we can always take this plane to be the x, y plane so that $\vec{e}_\rho = [\cos\theta(t), \sin\theta(t), 0]$. Then equation (II.4) describing the reflected field becomes, after gross-range compensation,

$$\begin{aligned} E(\nu_0 + \nu, t) &= \int f(\vec{r}) e^{j2\pi 2(\nu_0 + \nu)\vec{e}_\rho(t) \cdot \vec{r}} d^3r \\ &= \int \int \int f(x, y, z) e^{j2\pi 2(\nu_0 + \nu)[x\cos\theta(t) + y\sin\theta(t)]} dx dy dz \end{aligned}$$

$$= \int \int f(x, y) e^{j2\pi 2(\nu_0 + \nu)[x \cos \theta(t) + y \sin \theta(t)]} dx dy \quad (II.6)$$

where

$$f(x, y) \equiv \int f(x, y, z) dz.$$

Motion limited to a single plane results in a lack of information about the target in the direction perpendicular to that plane; the radar "sees" a two-dimensional target $f(x, y)$ which is the projection of the real three-dimensional target onto the plane of motion. We will assume that during the time the radar observes the target the following small angle approximations are valid,

$$\cos \theta(t) \approx 1$$

$$\sin \theta(t) \approx \theta(t)$$

$$2\nu y \theta(t)/c \ll 1.$$

Using these in (II.6) results in

$$E(\nu_0 + \nu, t) = \int \int f(x, y) e^{j2\pi 2[(\nu_0 + \nu)x/c + \nu_0 y \theta(t)/c]} dx dy.$$

Processing this recorded field into an image proceeds in two steps.

First we resolve in the x or *range* direction by inverse Fourier transforming over the band of transmitted frequencies. We obtain

$$\begin{aligned} & \int_{-\beta/2}^{\beta/2} E(\nu_0 + \nu, t) e^{-j2\pi 2(\nu_0 + \nu)x/c} d\nu \\ &= \int_{-\beta/2}^{\beta/2} \int \int f(\xi, y) e^{j2\pi 2[(\nu_0 + \nu)(\xi - x)/c + \nu_0 y \theta(t)/c]} d\xi dy d\nu \\ &= \int \left(\beta \int f(\xi, y) e^{-j2\pi 2\nu_0(x - \xi)/c} \operatorname{sinc} \left(\frac{x - \xi}{c/2\beta} \right) d\xi \right) e^{j2\pi 2\nu_0 y \theta(t)/c} dy \\ &\approx \int f(x, y) e^{j2\pi 2\nu_0 y \theta(t)/c} dy = \hat{f}_y(x, 2\nu_0 \theta(t)/c) \equiv \hat{f}_y(x, \phi(t)) \end{aligned}$$

where for convenience we define $\phi(t) = 2\nu_0 \theta(t)/c$. The function $\hat{f}_y(x, \phi(t))$ is the Fourier transform with respect to y of $f(x, y)$, evaluated at $\phi(t)$.

The second step is to achieve resolution in the y or *cross-range* direction. For each value of x the function we are left with, $\hat{f}_y(x, \phi(t))$, is the Fourier transform of the function we seek to compute, $f(x, y)$, but which has been distorted by a coordinate transformation, $\phi(t)$, due to the target's aspect history. In general, therefore, aspect compensation is required before imaging (inverse Fourier transforming in y) is possible. If $\phi(t)$ was known then aspect compensation could be performed by the inverse coordinate transformation

$$\int \hat{f}_y(x, \phi(t)) \delta(\phi(t) - v) \frac{d\phi(t)}{dt} dt = \int \hat{f}_y(x, \phi) \delta(\phi - v) d\phi = \hat{f}_y(x, v).$$

Following this with an inverse Fourier transform

$$\int \hat{f}_y(x, v) e^{-j2\pi y v} dv = f(x, y)$$

would resolve the target in y for this value of x . Repeating this process for each value of x would produce the full two-dimensional image. The coordinate transformation and Fourier transform can be combined into a single linear transformation

$$\begin{aligned} f(x, y) &= \int \hat{f}_y(x, v) e^{-j2\pi y v} dv = \int \left(\int \hat{f}_y(x, \phi(t)) \delta(\phi(t) - v) \frac{d\phi(t)}{dt} dt \right) e^{-j2\pi y v} dv \\ &= \int \hat{f}_y(x, \phi(t)) \frac{d\phi(t)}{dt} e^{-j2\pi y \phi(t)} dt. \end{aligned} \quad (II.7)$$

The problem in ISAR imaging is that $\phi(t)$ is unknown and therefore this transformation is unknown *a priori*.

II.5 Effects of Neglecting Aspect Compensation

The problem of aspect compensation is generally avoided in ISAR imaging by *assuming* that the target moved in such a smooth manner that $\phi(t) = \Omega t$ for some unknown constant Ω . In general $\phi(t)$ will actually be of the form $\phi(t) = \Omega t + \epsilon(t)$ where Ωt is the linear part of $\phi(t)$ and $\epsilon(t)$ is the nonlinear part (due to, for instance, target acceleration). If we take $\phi_e(t) = t$ as an estimate of the true $\phi(t)$ and use this to process the data via equation (II.7) the result is an estimate of $f(x, y)$ given by

$$\begin{aligned} f_e(x, y) &= \int \hat{f}_y(x, \phi(t)) e^{-j2\pi y t} dt = \int \left(\int f(x, \eta) e^{j2\pi \eta \phi(t)} d\eta \right) e^{-j2\pi y t} dt \\ &= \int f(x, \eta) \left(\int e^{j2\pi \eta \epsilon(t)} e^{-j2\pi t(y - \Omega \eta)} dt \right) d\eta \\ &= \int f(x, \eta) h(y, \eta) d\eta \end{aligned}$$

where

$$h(y, \eta) = \int e^{j2\pi \eta \epsilon(t)} e^{-j2\pi t(y - \Omega \eta)} dt, \quad (II.8)$$

represents the cross-range impulse response of the ISAR imaging system. That is $h(y, \eta)$ is the cross-range image of a point scatterer at position η on the target. If the target actually moved so that $\phi(t) = \Omega t$ then $h(y, \eta) \propto \text{sinc} \Delta t (y - \Omega \eta)$, where Δt is the length of time the radar observed the target, and so $f_e(x, y) \approx f(x, y/\Omega)$ and this approach works. Diffraction limited resolution is achieved although there is an unknown scaling in y due to the unknown constant Ω .

On the other hand if there is any non-linearity in the aspect history (i.e. $\epsilon(t) \neq 0$) then $h(y, \eta)$ will have a width greater than that of the diffraction limit. If $|\eta\epsilon(t)|$ is small then (II.8) can be approximated by

$$h(y, \eta) \approx \int (1 + j2\pi\eta\epsilon(t)) e^{-j2\pi t(y - \Omega\eta)} dt$$

$$\propto [\text{sinc } \Delta t z + 2\pi\eta\hat{\epsilon}(z) * \text{sinc } \Delta t z]_{z=y-\Omega\eta}$$

where $\hat{\epsilon}$ is the Fourier transform of ϵ . The first term is the diffraction limited impulse response. The second term produces blurring beyond the diffraction limit. It grows stronger with larger $|\eta|$ and has a width beyond the diffraction limit equal to the spectral width of $\epsilon(t)$. For large $|\eta\epsilon|$ the width of $h(y, \eta)$ is approximately $2|\eta| \cdot |d\epsilon/dt|_{\max}$ (twice the maximum instantaneous frequency of $\exp[j2\pi\eta\epsilon(t)]$) beyond the diffraction limit. In either case the impulse response is blurred beyond the diffraction limit when there is any non-linearity in the aspect history with the blurring being worst at the edges (largest $|\eta|$) of the target.

II.6 Learning Aspect Compensation

Computing $f(x, y)$ from $\hat{f}_y(x, \phi(t))$ is an underdetermined problem because we lack necessary information, namely $\phi(t)$, needed to obtain the correct solution. For any guess we might make at $\phi(t)$, say $\phi_e(t)$, we can produce an image, call it $f_e(x, y)$, through the transformation described in equation (II.7). However $f_e(x, y)$ will faithfully represent $f(x, y)$ only if we somehow get $\phi_e(t) \approx \phi(t)$. What we need is some additional constraint on the problem which when satisfied forces $\phi_e(t) \approx \phi(t)$.

Neglecting aspect compensation as discussed above (which is the common practice in ISAR imaging) essentially puts a constraint on $\phi(t)$; in order to produce a well-focused image it is required that $\phi(t) = \Omega t$. The problem with this approach is that we don't have any control over $\phi(t)$, since we don't control the target's motion, and hence we can't enforce this constraint. A better approach, we think, is to put a constraint on something we do have control over, namely the image. We have some control over the image by our choice of an aspect history estimate, $\phi_e(t)$. The obvious constraint is to require that the image be as well-focused as possible. Therefore we loosen constraints on the target's motion by adopting a parametric form for $\phi_e(t)$ with degrees of freedom, such as

$$\phi_e(t; a_1, \dots, a_n) = a_1 t + a_2 t^2 + \dots + a_n t^n,$$

while at the same time seeking to enforce the constraint that the image obtained by processing with $\phi_e(t)$ be focused. Below we describe a simple quantitative measure which we can use to enforce this constraint.

With $\epsilon(t; a_1, \dots, a_n) \equiv \phi_e(t; a_1, \dots, a_n) - \phi(t)$ denoting the error in the estimate of the aspect history we get an image of the target through the linear transformation in (II.7) (we will no longer explicitly show the dependence of ϕ_e and ϵ on a_1, \dots, a_n)

$$f_e(x, y; a_1, \dots, a_n) = \int \hat{f}_y(x, \phi(t)) \frac{d\phi_e(t)}{dt} e^{-j2\pi y\phi_e(t)} dt$$

$$\begin{aligned}
&= \int \left(\int f(x, \eta) e^{j2\pi\eta[\phi_e(t) - \epsilon(t)]} d\eta \right) \frac{d\phi_e(t)}{dt} e^{-j2\pi y \phi_e(t)} dt \\
&= \int f(x, \eta) \left(\int e^{-j2\pi\eta\epsilon} e^{-j2\pi(y-\eta)\phi_e} d\phi_e \right) d\eta \\
&= \int f(x, \eta) h(y, \eta; a_1, \dots, a_n) d\eta
\end{aligned}$$

where

$$h(y, \eta; a_1, \dots, a_n) = \int e^{-j2\pi\eta\epsilon} e^{-j2\pi(y-\eta)\phi_e} d\phi_e$$

is the impulse response of our imaging system. The impulse response depends on a_1, \dots, a_n because $\phi_e(t)$ and $\epsilon(t)$ do. It is easy to show that

$$f_e(x, y; \alpha a_1, \dots, \alpha a_n) = \alpha f_e(x, \alpha y; a_1, \dots, a_n)$$

so scaling $\phi_e(t)$ by a constant factor scales the image in the y dimension but has no effect on how well focused the image is. Assuming we have no *a priori* information about the target which would allow us to choose the proper scaling we will always scale the parameters a_1, \dots, a_n so that $\Delta\phi_e = 1$, where $\Delta\phi_e \equiv \max[\phi_e(t)] - \min[\phi_e(t)]$. This insures that the scale of the image remains constant.

By an analysis of the impulse response similar to that used in section II.5 above we can show that image blurring increases as ϵ grows larger and that for a given $\epsilon(t)$ the blurring is most severe at the edges of the image (large $|\eta|$). Since the width of a blurred image is the width of the well-focused image plus the width of the blurring at the edges we are led to construct an "energy" function $w(x; a_1, \dots, a_n)$, defined by

$$\int_{-w(x; a_1, \dots, a_n)/2}^{w(x; a_1, \dots, a_n)/2} |f_e(x, y; a_1, \dots, a_n)|^2 dy = c \int_{-\infty}^{\infty} |f_e(x, y; a_1, \dots, a_n)|^2 dy,$$

as a measure of the width of our image and hence the amount of blurring. For a given x , w is the number such that a fraction c , $0 < c < 1$, of the total energy of this strip of the image lies in the region $x \in [-w/2, w/2]$; we take $c < 1$ because due to noise and the sidelobes of the impulse response all of the energy cannot be confined to a finite region but only "most" of the energy. Our system can learn the necessary aspect compensation by searching its state space (the parameters a_1, \dots, a_n) until it minimizes this energy function and hence the image blurring. At this point the output of the system is a well-focused strip of the image (provided there are enough degrees of freedom in $\phi_e(t)$ to allow a good approximation to $\phi(t)$). The constant scale constraint $\Delta\phi_e = 1$ is essential because it insures that changes in w are due to changes in the amount of blurring in the image and not merely to changes in the scale of the image. Since blurring is largest for largest $|\eta|$ aspect compensation can be learned most accurately by minimizing $w(x_0; a_1, \dots, a_n)$ where $w(x_0; a_1, \dots, a_n) \geq w(x; a_1, \dots, a_n)$ for all $x \neq x_0$, that is by learning on that strip of the image which is widest. The remainder of the image can then be computed strip

by strip through (II.7), the kernel of which our system has just learned by minimizing $w(x_0; a_1, \dots, a_n)$.

To summarize, the algorithm is as follows. Initialize the parameters (we take $a_1 = 1, a_{i \neq 1} = 0$). Find x_0 such that $w(x_0; a_1, \dots, a_n) \geq w(x; a_1, \dots, a_n)$. Then minimize $w(x_0; a_1, \dots, a_n)$ with respect to a_1, \dots, a_n producing the one strip of the image, $f(x_0, y)$. Finally use the aspect history estimate $\phi_e(t; a_1, \dots, a_n)$ learned in the last step to compute the rest of the image, $f(x \neq x_0, y)$, strip by strip.

Figures II.1a-II.1c show the results of a simulation of this imaging technique. We considered a target consisting of a collection of point scatterers arranged in a "V" shape (figure II.1a) and simulated the radar returns corresponding to an aspect history of $\phi(t) = t^2, 0 \leq t \leq 1$. Figure II.1b shows the image produced by the standard processing approach which takes $\phi_e(t) = t$. We allowed our processor 2 degrees of freedom, $\phi_e(t) = at + bt^2$, which together with the scaling constraint, $\Delta\phi_e = 1$ for $0 \leq t \leq 1 \Rightarrow a + b = 1$, defined a one-dimensional state space. By searching over this state space until $w(x_0; a, b)$ was minimized (we took $c = 0.95$) our processor produced the image shown in figure II.1c.

We also applied this technique in a simple acoustic imaging experiment. Figure II.2a is a picture of the target, a collection of ping-pong balls arranged in a "V" shape (each row of balls is about 500mm long). While illuminating the target with acoustic pulses and recording the reflected field we varied its position so that it had an aspect history $\theta(t) = 2.5^\circ(t + t^2), 0 \leq t \leq 1$. We then produced images using both the standard processing approach and our algorithm (with $\phi_e(t) = at + bt^2$ and $a + b = 1$ as above). The resulting images are displayed in figures II.2b and II.2c respectively.

II.7 Optical Implementation

The computation required to produce one strip of the image, given in equation (II.7), is a shift-variant linear transformation which optics is well suited to implement. Figure II.3 shows an optical implementation of the iterative ISAR algorithm we described above. It consists of a programmable optical matrix-vector multiplier with an electronic feedback path from the output to the matrix. Given a matrix

$$K(t, y; a_1, \dots, a_n) = \frac{d\phi_e(t; a_1, \dots, a_n)}{dt} e^{-j2\pi y \phi_e(t; a_1, \dots, a_n)}$$

the optical section computes the linear transformation

$$f_e(x_0, y; a_1, \dots, a_n) = \int f(x_0, \phi(t)) \frac{d\phi_e(t; a_1, \dots, a_n)}{dt} e^{-j2\pi y \phi_e(t; a_1, \dots, a_n)} dt.$$

The feedback circuitry varies the matrix at each iteration through the parameters a_1, \dots, a_n and computes the energy $w(x_0; a_1, \dots, a_n)$ from the output $f_e(x_0, y; a_1, \dots, a_n)$. It continues to iterate until it finds the global energy minimum. At this point the output is a well

focused image of a strip of the target around x_0 and the processor has learned the necessary aspect compensation. With the matrix fixed the remaining data, $\hat{f}_y(x, \phi(t))x \neq x_0$, is then fed through the processor allowing the full two-dimensional image to be produced strip by strip.

We set up a preliminary version of this processor using a magneto-optic device (MOD) for the programmable matrix, a 1D CCD array to detect the output $|f_e(x, y; a_1, \dots, a_n)|^2$, and a PC to update the matrix. We did not implement the first half of the optical matrix-vector multiplier, which spreads $f(x, \phi(t))$ in the y direction to form the product $f(x, \phi(t))K(t, y; a_1, \dots, a_n)$ in the plane of the matrix. Instead we formed this product digitally and fed the result into the MOD.

Figures II.4a-II.4c show the output of the optical processor in various stages of learning to image a strip of the "V" shaped simulated target discussed above. The input data was the same as that used in the simulation above. The initial output (with $\phi_e(t; a_1, \dots, a_n) = t$) is depicted in figure II.4a. Here the error in the aspect history estimate is large and the image is quite blurred. From this point the processor starts moving through its state space, (a_1, a_2) , minimizing the energy $w(x_0; a_1, a_2)$. Figure II.4b shows the output at an intermediate stage. Finally the processor settles at the output shown in figure II.4c which had the minimum width and in which the two point scatterers in this strip of the image are well-focused.

References

- [1] Wehner, D.R., *High Resolution Radar*, Artech House, 1987
- [2] Chen, C.C., Andrews, H.C., "Target-Motion-Induced Radar Imaging", *IEEE Trans. Aeosp. Electron. Syst.*, vol AES-16
- [3] Farhat, N., Chu, T.H., "Tomography and Inverse Scattering", ICO-13

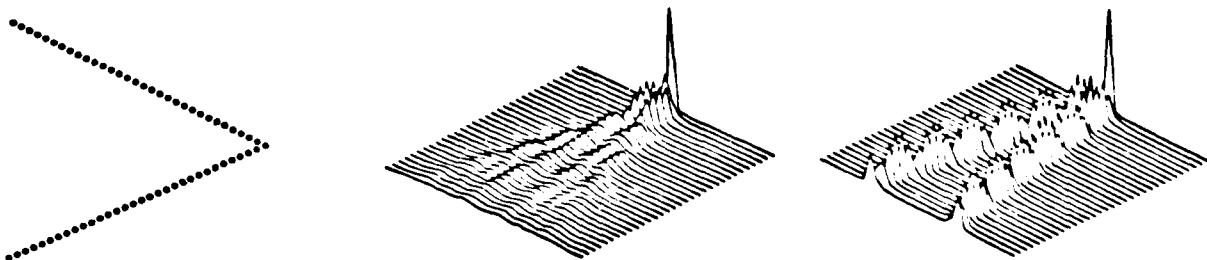


Fig. II.1a, II.1b, II.1c

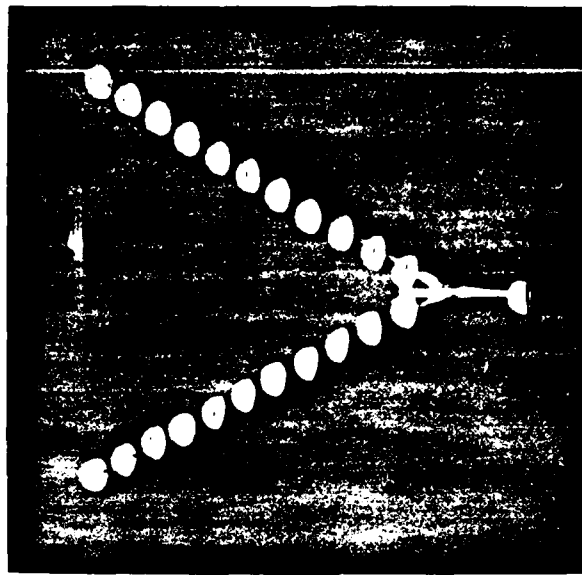


Fig. II.2a

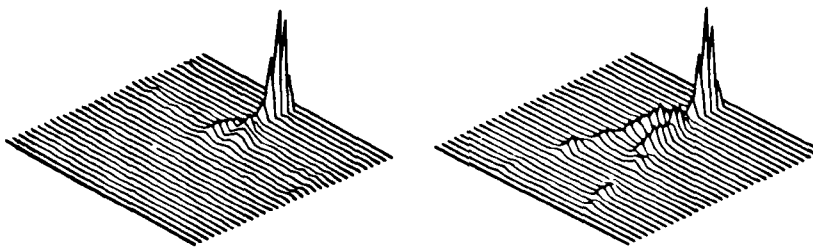


Fig. II.2b, II.2c

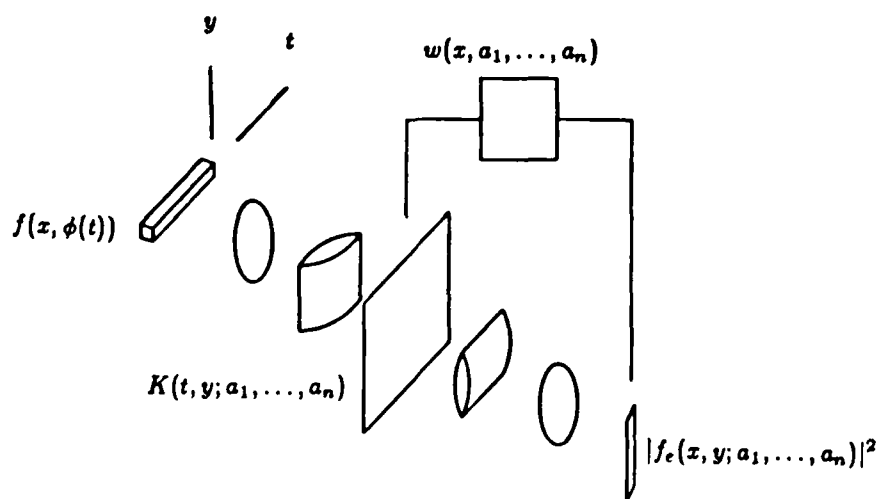


Fig. II.3

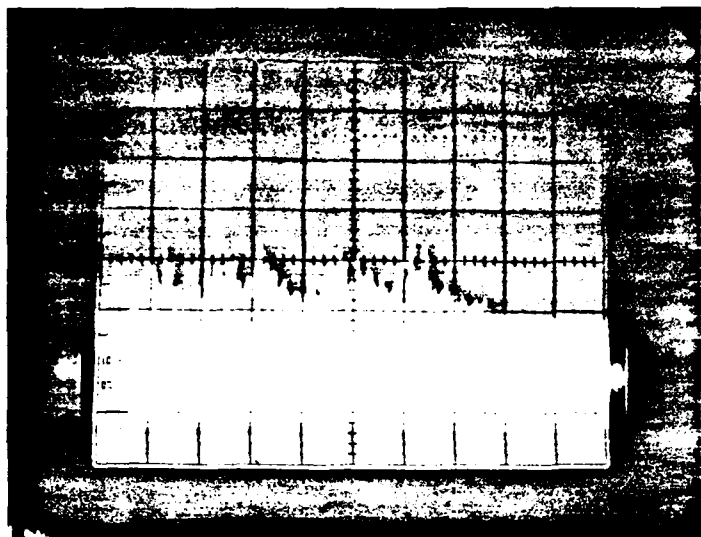


Fig. II.4a

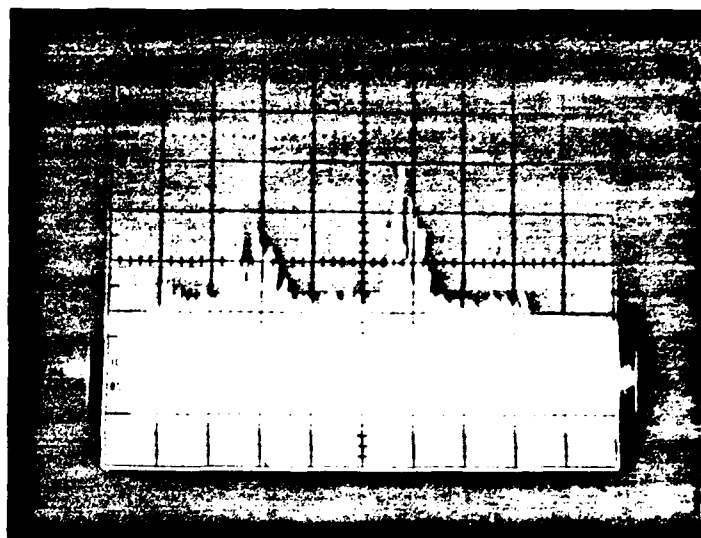


Fig. II.4b

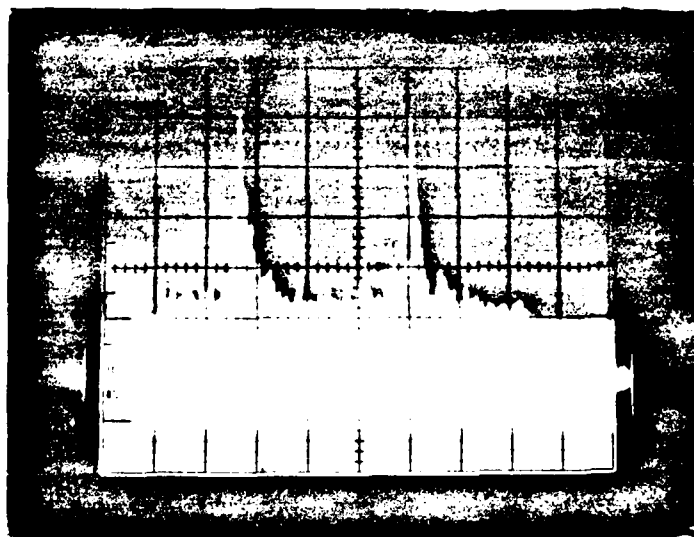


Fig. II.4c

III

CAPACITY OF OPTICAL CORRELATORS

III.1 Introduction

Vander Lugt correlators have been used for a long time in optical pattern recognition[1]. In the typical implementation, shown in figure III.1, the Fourier transform of an input image is used to read out a hologram containing the Fourier transform of a reference image. This diffracted beam is then inverse Fourier transformed to produce the correlation between the input and reference images on the output plane. Vander Lugt correlators are typically used as pattern recognition systems. Whether or not a peak is present at the output of the correlator determines whether or not the input image is sufficiently close to the stored reference. Recently much work has been done on the use of a Vander Lugt correlator to pattern classification [2][3]. In this case, the correlator distinguishes whether the input is a member of one of two classes with each class being composed of many images. Typically, a reference filter is formed as a linear combination of the images in both classes and the presence or absence of a peak at the correlation plane determines which class the input belongs to. In this paper, we discuss the capacity of the Vander Lugt correlator. This is to say we estimate the maximum number of images that can be stored in the reference filter before the system begins to misclassify images. This capacity has been studied in great detail for systems without shift invariance (e.g. perceptrons). The classic results from pattern recognition about the capacity of a linear discriminant function do not directly apply in this case because the VanderLugt correlator is shift invariant. In this paper, we will discuss the capacity of the system incorporating the shift invariance of the Vander Lugt correlator. We will also discuss the effect on the capacity of binarizing the reference filter and lastly we will demonstrate that by using a volume hologram to record the filter, the capacity of the system is greatly increased, as well as becoming capable of multi-class classification.

III.2 Capacity of Linear Filters

In the most common pattern classification scheme, the inner product is performed between the input image $\phi(x, y)$ composed of N pixels and a reference filter $h(x, y)$.

$$O = \sum_{y'=1}^{\sqrt{N}} \sum_{x'=1}^{\sqrt{N}} h(x', y') \phi_i(x', y') \quad (III.1)$$

Comparing the output O with a pre-set threshold, determines which of the two classes the input belonged to. A standard method of forming the reference filter is as a linear

combination of the images in both classes:

$$h(x, y) = \sum_{i=1}^M w_i \phi_i(x, y) \quad (III.2)$$

The weights w_i can be chosen through a variety of training algorithms such as the perceptron learning algorithm. It is a well known result that the capacity of such a system is [4]

$$M = 2N \quad (III.3)$$

where N is the number of pixels in each image. In this paper, we will consider the construction of a simpler filter in which the weights are binary.

$$w_i = \begin{cases} 1 & \text{if } \phi_i \in \text{Class I} \\ 0 & \text{if } \phi_i \in \text{Class III} \end{cases} \quad (III.4)$$

In other words, the filter is formed by simply summing the images belonging to class 1, while ignoring those in class 2. This is implemented in a Vander Lugt correlator, by multiply exposing the hologram to the images in class 1 while doing nothing for the images in class 2. Classification can then be performed by detecting and thresholding the output at the center of the correlation plane. For the remainder of the paper, we will assume that the images $\phi(x, y)$ consist of binary N pixels, each pixel being a bipolar (ie 1 or -1), independent random variable. Under these assumptions, the capacity of the VanderLugt correlator using the peak-only detection scheme can be found by solving the following transcendental equation[5]:

$$M = \frac{N}{4 \log(M^3/N)} \quad (III.5)$$

As $N \rightarrow \infty$, the above expression asymptotically approaches

$$M = \frac{N}{8 \log N} \quad (III.6)$$

Thus the use of the simpler method for constructing the reference filter, results in a relatively modest loss in capacity by a factor of $16 \log N$

III.3 Capacity of Shift Invariant Filters

Because Vander Lugt correlators are inherently shift invariant it is possible to classify prescribed images and their shifted versions as well. In order to implement a shift invariant classification scheme, detection at the output is done over the entire correlation plane. As a result the detection of a peak anywhere in the output plane determines whether the input is a member of class 1 or a shifted form of a member in class 1. Figure III.2a shows a cross section through the origin of the digital correlation of an input image with a

filter containing only one image. The resulting output shows a single correlation peak and relatively small sidelobes. When the reference is constructed by adding 3 images (figure III.2b, the sidelobe structures shows a significant rise in amplitude. However, since only the single correlation peak lies above the threshold, classification of the input image is still performed correctly. However, when the number of reference images is increase to 6 (Fig III.2c), there are now two peaks which lie above the threshold level. As a result, the system can no longer decide whether the input image is a member of class 1 or a shifted version of a member of class 1. Therefore, we expect that the capacity of the shift invariant system is smaller. For the relatively simple method of filter construction , we can readily derive an analytic capacity for the shift invariant correlator. In the shift invariant case, the Vander Lugt system performs a correlation between one of the input image $\phi(x, y)$ and the reference filter $h(x, y)$

$$O(x, y) = \sum_{y'=1}^{\sqrt{N}} \sum_{x'=1}^{\sqrt{N}} h(x', y') \phi_i(x' + x, y' + y). \quad (III.7)$$

For the case where the filter is constructed by simply summing the images in class 1 (multiple exposure) and assuming the same input statistics for each image, the capacity of the shift invariant Vander Lugt system is given by the solution of the following transcendental equation [5]

$$M = \frac{N}{4 \log(M^3 N)} \quad (III.8)$$

Asymptotically, the capacity approaches

$$M = \frac{N}{16 \log N} \quad (III.9)$$

Thus, the capacity is decreased by only a factor of two from that of the non shift invariant system. This result is important since there is to our knowledge no prior estimate for the loss in capacity due to shift invariance. For the case considered here (the filter derived as a simple sum), the loss is very small; a factor of 2. To verify the theoretical capacity of the correlator, 100 computer trials were averaged to determine the capacity for various N . For each trial, two random vectors were generated to form the initial reference filter. Each image was correlated to determine whether classification was performed correctly. If no error occurred, a new random image was added to the reference filter and correlation with all the image was done. The number of images in the reference was increased until a misclassification occurred. At this point, the capacity was said to be one less than the number of images stored in the reference.

Figure III.3 shows the capacity of both the peak only and shift invariant systems as a function of the number of pixels, N , in the image. Experimental simulations show good agreement with theoretical predictions. It is important to note that because the simulations were performed in the regime of small N , the transcendental equations for the capacity (eqs (III.5) and (III.8)) were used to plot the theoretical curves.

III.4 Capacity of Binary Filters

As demonstrated above, the capacity of the VanderLugt system can be very large. One potential limitation that might prevent us from actually implementing such a large pattern classification system is the accuracy with which the hologram can record the reference. To get a feel for the susceptibility of the system to nonlinearities and inaccuracies, we considered the capacity of the Vander Lugt correlator when the reference filter has been binarized.

In this case, the reference filter consists of a thresholded version of the filter generated from the multiple exposure algorithm

$$h(x, y) = \text{sgn} \left[\sum_{i=1}^M w_i \phi_i(x, y) \right] \quad (\text{III.10})$$

Again, assuming that the input pixels consists of bipolar independent random variables, we find that the capacity of the binary Vander Lugt correlator is asymptotically

$$M = \frac{N}{8\pi \log N} \quad (\text{III.11})$$

There is only a further $\pi/2$ reduction in capacity from that of the non binarized shift invariant filter. In figure III.4, a comparison is made between the linear and a binary filters. In both cases, the input images had 256 pixels and the reference filter contained 3 images. As seen from the figure, the sidelobe level of the binary correlator is significantly larger than that for the linear filter. As a result, as additional images are added to the reference filter, the binary correlator will begin to misclassify sooner. This will correspondingly lead to a lower capacity (theoretical and experimental) for the binary correlator. In figure III.5, the capacity of both the binarized and nonbinarized filters are plotted as a function of the number of pixels in the image. Again, computer simulations demonstrate a good agreement with theoretical predictions.

III.5 Capacity of the Volume VanderLugt Correlator

In this section, we consider the use of a volume hologram to record the reference filters in a Vander Lugt correlator[6]. We expect that because information is recorded in three dimensions as opposed to the two dimensions for plane holograms, the storage capacity of the volume VanderLugt correlator is increased. Let us first consider how a volume Vander Lugt correlator operates (Fig III.6). Consider the correlation between two point sources. In the recording stage (Fig III.6a), the point source generates a plane wave which interferes with a reference wave to form a grating which is recorded in the volume hologram. When an input point source at the same position is presented to the correlator (Fig III.6b), a new plane wave reads out the stored grating. The diffracted plane wave is then focussed

to form the expected correlation peak at the output. If, however, the input point source is shifted in the direction parallel to the plane of incidence (Fig III.6c), the plane wave that is generated will not be Bragg matched with the grating in the volume hologram. Consequently no diffracted wave will be produced and no correlation spot will be formed. In the direction perpendicular to the plane of incidence, the volume hologram exhibits very little Bragg sensitivity and a correlation can still be read out. As a result, shifts of the input in a direction parallel to plane of incidence will not be recognized, while in the perpendicular direction the correlator remains shift invariant.

For an arbitrary input, $A(x, y)$ and reference image, $R(x, y)$, it can be shown that the output of the volume Vander Lugt correlator is [7]

$$O(x, y) = [A(x, y) * R(x, y)] \text{sinc}(\alpha x) \quad (III.12)$$

where $\alpha = T \sin \theta / 2\lambda F$ and $*$ is the correlation operator. T is the thickness of the hologram, θ is the Bragg angle, and F is the focal length of the inverse Fourier transform lens. In other words, the output of the correlator consists of the correlation between the input and reference apodized by a sinc function whose width is determined by the thickness of the volume hologram.

To experimentally demonstrate this apodizing effect, the auto correlation of an O was performed using the volume Vander Lugt correlator. Figure III.7 shows a digitally generated autocorrelation of an O which simulates a standard Vander Lugt correlator with a reference O recorded on a plane hologram. In the volume Vander Lugt correlator, the reference O was recorded on a lithium niobate crystal measuring 25x25x5mm. The reference beam was situated such that the plane of incidence was in the horizontal direction. Figure III.8a shows the output of the volume Vander Lugt correlator when the input O is positioned at the same plane as the reference O . The output consists of the standard correlation of the two O 's multiplied by the horizontal sinc function. When the input O is shifted in the direction parallel to plane of incidence (Fig III.8b), the correlation shifts and only correlation structure to one side of the peak is presented at the output. The smaller spot lying to the right of the primary horizontal band corresponds to the very strong correlation peak lying in the first sidelobe of the apodizing sinc function. Further shifts of the input as shown in figure III.8c, merely reads out the correlation structure further from the peak.

The Bragg selectivity in the volume VanderLugt correlator allows one to perform multi-class categorization of the input images[8]. In the recording stage (Fig III.9a), a set or reference filters is recorded by interfering each with a reference beam separated by the angular bandwidth of the volume hologram. When an input image is presented to the volume Vander Lugt correlator (Fig III.9b), a set of correlations is performed simultaneously and presented spatially distributed at the output. The Bragg selectivity of the hologram guarantees that the correlation bands will not interfere with each other. As a result, by detecting which band the correlation peak appear, determines which of many classes the input image belongs to.

We can consider each correlation band as a separate output channel performing a simple pattern classification task independent of the other channels. By assuming the

same input statistics for the images in each class, the capacity of each output channel can be analytically derived. In this case, the maximum number of images that can be stored was found to be equal to that of the standard VanderLugt correlator (eq III.5). Asymptotically, the capacity of each channel approaches

$$M' = \frac{N}{16 \log N} \quad N \rightarrow \infty \quad (III.13)$$

The number of output channels, K , that can be stored in the volume hologram is

$$K = \frac{TL \sin \theta}{\lambda F} \quad (1 \leq K \leq N) \quad (III.14)$$

where L is the actual dimension of the output detector array in the direction parallel to the plane of incidence. Hence the total capacity of the system is

$$M_{Total} = \frac{KN}{16 \log N} \quad (N \rightarrow \infty; 1 \leq K \leq N) \quad (III.15)$$

Thus the effect of the using a volume hologram is that the capacity is increased by the number of output channels the hologram can support and one to perform multi-class classification. However, one drawback is the partial loss of shift invariance in one direction that results from the use of a volume hologram.

III.6 Conclusion

In conclusion, we have demonstrated that the capacity of a Vander Lugt correlator without shift invariance is $N/8 \log N$ for the simple additive filter. By incorporating the shift invariance inherent in an optical correlator, the capacity is only decreased by a factor 2. Furthermore, by binarizing the reference filter, there is a further loss by a factor of $\pi/2$. However, by utilizing a volume hologram to record the reference filter, the capacity of the correlator is increased by a factor that can be as high as N with a proportional loss in shift invariance.

References

- [1] A. VanderLugt, *IEEE Trans. Inf. Th.*, IT-10 139 (1964)
- [2] D. Casasent *Appl. Opt.*, 23 1620 (1984)
- [3] A. Lohmann and C. Thum *JOSA*, 23 1503 (1984)
- [4] F. Rossblatt *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms* (Spartan Books, Washington D.C. 1961)

- [5] J. Yu, F. Mok, and D. Psaltis *submitted to Applied Optics*
- [6] S.I. Stepanov and V. D. Gural'nik *Sov. Tech. Phys. Lett.* ,8 49 (1982)
- [7] J. Yu and D. Psaltis *To be published*
- [8] C. Guest and C. Gaylord *Appl. Opt.*,19 1201 (1983)

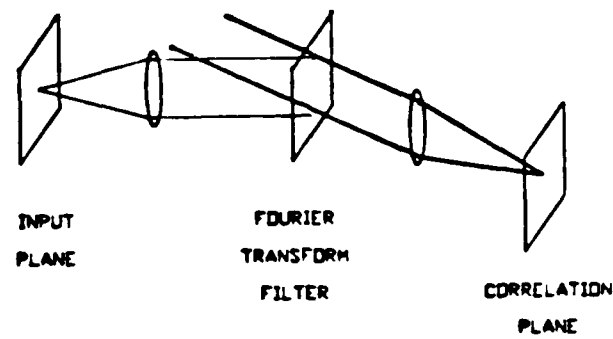


Fig. III.1. Vander Lugt correlator.

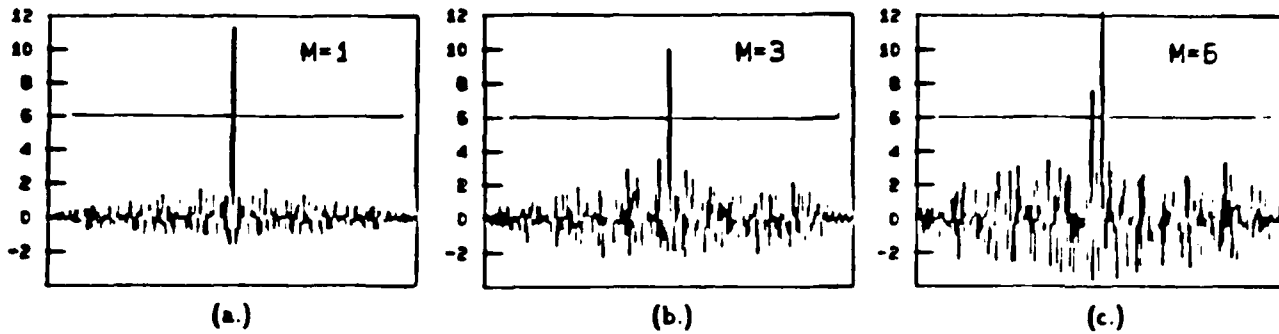


Fig. III.2. Digital correlations of a shift invariant filter.

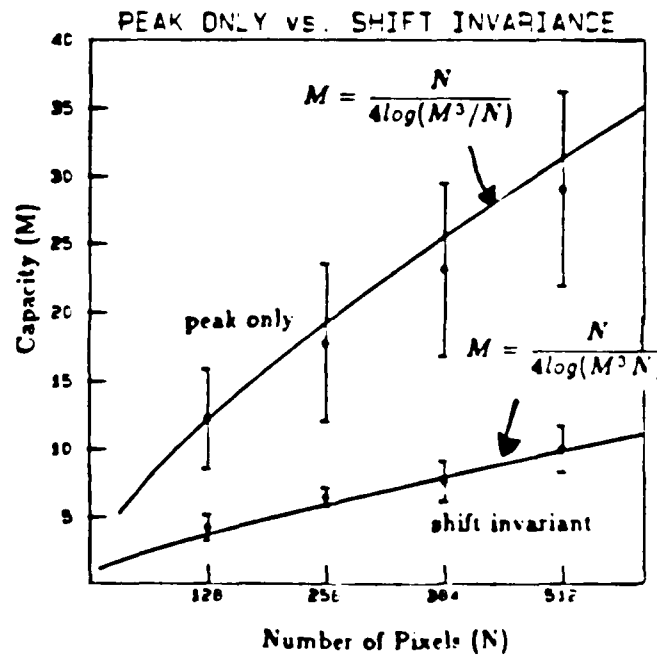


Fig. III.3. Capacity curves of the peak only and shift invariant filters.

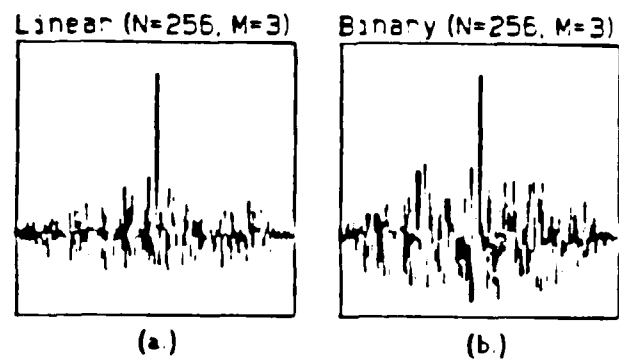


Fig. III.4. Digital correlations of linear and binary filters.

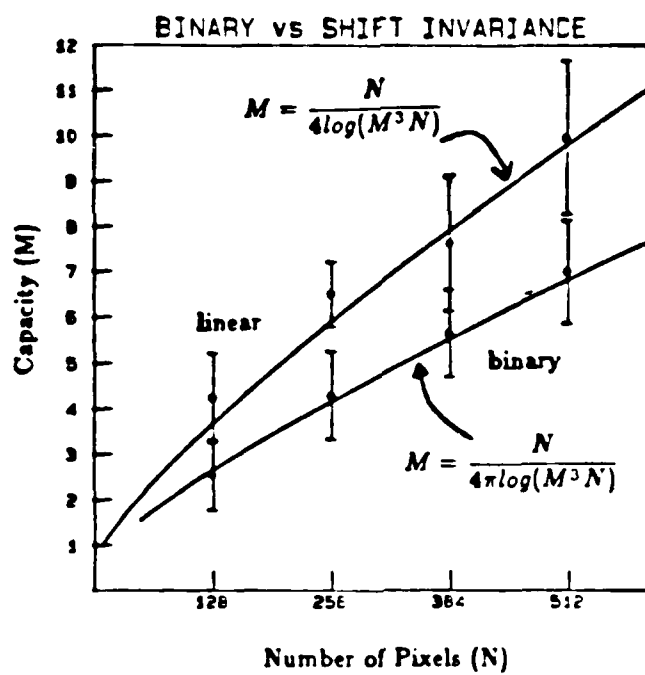


Fig. III.5. Capacity curves of shift invariant and binary filters.

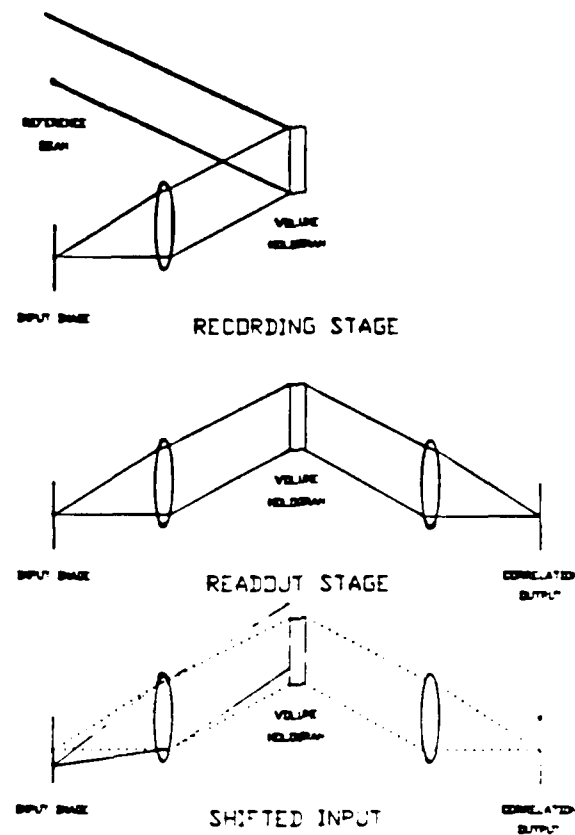


Fig. III.6. Recording and readout of a volume hologram.

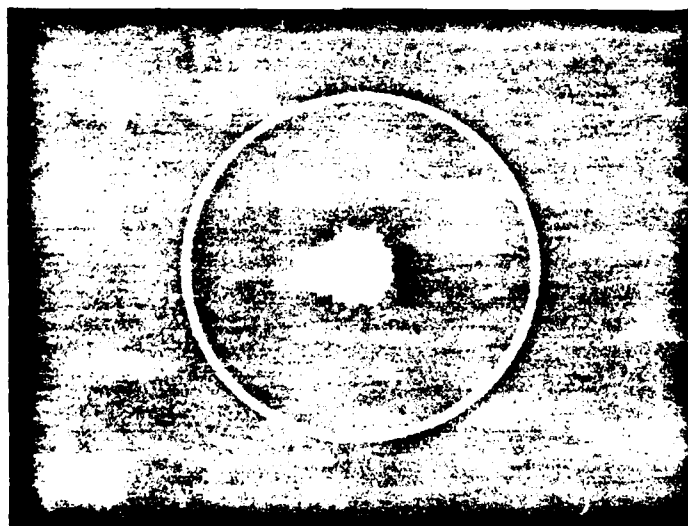


Fig. III.7. Digal autocorrelation of an "O".

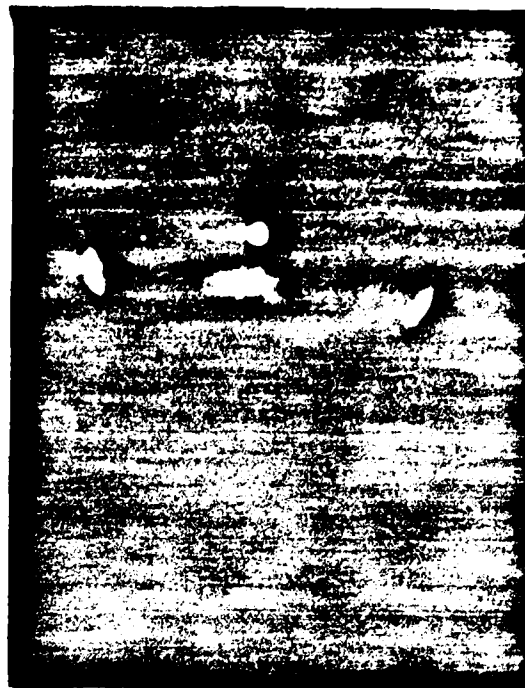
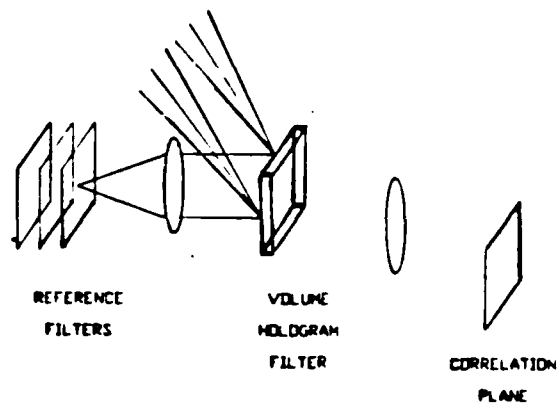
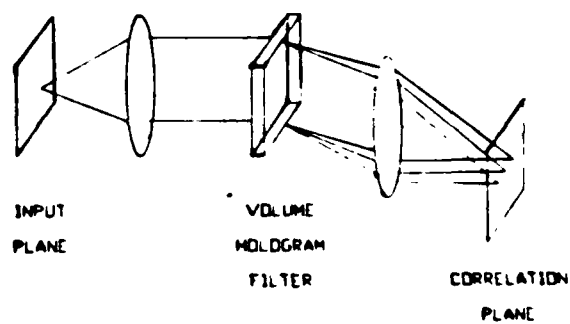


Fig. III.8. Experimental outputs of volume Vander Lugt correlator



RECORDING STAGE

(a.)



READOUT STAGE

(b.)

Fig. III.9. Recording and readout stage of the multi-class. categorization volume Vander Lugt correlator.

IV

PHOTOREFRACTIVE INTEGRATED OPTICAL VECTOR MATRIX MULTIPLIER

IV.1 Introduction

Vector matrix multiplication is central to a wide range of signal processing systems, including neural networks, large scale interconnection networks, and linear transformations. Bulk optical implementations of vector matrix multiplication are well developed and quite successful [10]. The classical optical vector matrix multiplier uses a cylindrical lens to expand a one dimensional array of light sources across a two dimensional spatial light modulator on which the matrix is recorded. A second, orthogonal, cylindrical lens focuses light from the SLM onto an array of detectors. Each element of this array receives light from each element in the input vector, modulated by the appropriate matrix element. This method of implementing vector matrix multiplication has been applied in a number of signal processing architectures.

Despite the success of these devices, the implementation of vector matrix multipliers in integrated optics offers certain advantages which are difficult to achieve in bulk. Integrated devices are typically much smaller and lighter than their bulk counterparts. Fabrication of a large number of identical integrated devices is relatively easy due to the advanced state of photolithography. In addition, integration leaves the dimension out of the integrated plane free for optical control of the matrix interconnecting vectors in the guided plane. In bulk devices dynamic control of this matrix may be relatively difficult. In this paper we propose implementing a vector matrix multiplier in integrated optics using photorefractive holography in the volume of the waveguide. A schematic architecture for this device is shown in figure IV.1. A vector is input through a channel waveguide array integrated on a suitable substrate. Light from these channels is collimated by an integrated lens [7] before reading out a matrix of holographic gratings formed in a slab waveguiding region. The diffracted vector is refocused into an output channel waveguide array. This architecture performs vector matrix multiplication with as many degrees of freedom as the bulk architecture by utilizing the Bragg selectivity of volume holograms to substitute for one dimension of the bulk spatial light modulator.

IV.2 Matrix Implementation via Integrated Volume Holograms

Vector matrix multiplication using diffraction gratings may be described via coupled wave equations in Fourier space. The electric field vector $\vec{E}(\vec{r})$ may be described at any plane along the optical axis of the system in terms of a complete set of guided and unguided modes of the slab waveguide. Letting $A_\nu(\vec{k})$ represent the component of $\vec{E}(\vec{r})$ in the mode

ν at spatial frequency \vec{k} we may express $\vec{E}(\vec{r})$ as

$$\vec{E}(\vec{r}) = \int \int \int \sum_{\nu} A_{\nu}(\vec{k}) \vec{\mathcal{E}}_{\nu}(\vec{k}, \vec{r}) d\vec{k} \quad (IV.1)$$

$\vec{\mathcal{E}}_{\nu}(\vec{k}, \vec{r})$ is the field distribution for the mode ν at spatial frequency \vec{k} . The effect on $\vec{E}(\vec{r})$ of the hologram recorded in the volume of the slab may be described by substituting (IV.1) into the Helmholtz equation and applying the slowly varying envelope approximation. From this approach we obtain

$$\int \int \int \sum_{\nu} \frac{\partial A_{\nu}(\vec{k})}{\partial z} j k_{\nu z} \frac{\partial \vec{\mathcal{E}}_{\nu}(\vec{k}, \vec{r})}{\partial z} d\vec{k} = \sum_{\vec{K}_g} \Delta k^2(\vec{K}_g, \vec{r}) \int \int \int \sum_{\nu} A_{\nu}(\vec{k}) \vec{\mathcal{E}}_{\nu}(\vec{k}, \vec{r}) d\vec{k} \quad (IV.2)$$

where $\Delta k^2(\vec{K}_g, \vec{r})$ is proportional to the holographic perturbation to the dielectric constant at spatial frequency \vec{K}_g . By applying the orthogonality condition

$$\int \int \int \vec{\mathcal{E}}_{\nu}(\vec{k}, \vec{r}) \cdot \vec{\mathcal{E}}_{\nu'}^*(\vec{k}', \vec{r}) d\vec{r} = \delta_{\nu\nu'} \delta(\vec{k} - \vec{k}') \quad (IV.3)$$

Equation (IV.2) may be reduced to

$$\frac{\partial A_{\nu}(\vec{k}_1)}{\partial z} = -j \sum_{\nu} \sum_{\vec{K}_g} \kappa_{\nu \vec{K}_g} A_{\nu}(\vec{k}_2) \quad (IV.4)$$

where

$$\vec{k}_1 - \vec{k}_2 = \vec{K}_g \quad (IV.5)$$

and

$$\kappa_{\nu \vec{K}_g} = \frac{1}{k_{\nu z}} \int \int \int \Delta k^2(\vec{K}_g, \vec{r}) \vec{\mathcal{E}}_{\nu}(\vec{k}_p - \vec{K}_g, \vec{r}) \cdot \vec{\mathcal{E}}_{\nu_p}^*(\vec{k}_p, \vec{r}) d\vec{r} \quad (IV.6)$$

In deriving equation (IV.4) we have ignored the finite nature of the hologram. Equation (IV.4) may be expressed

$$\frac{\partial \vec{A}}{\partial z} = \vec{\bar{H}} \vec{A} \quad (IV.7)$$

Each $A_{\nu}(\vec{k}_1)$ with \vec{k}_1 satisfying equation (IV.5) forms a component of \vec{A} . The components $\vec{\bar{H}}$ are given by

$$\vec{\bar{H}}(\vec{k}_1, \vec{k}_2) = \kappa_{\nu(\vec{k}_1 - \vec{k}_2)} \quad (IV.8)$$

The solution to equation (IV.7) is $\vec{A} = \exp(\vec{\bar{H}}z) \vec{A}_0$. If all of the eigenvalues of $\vec{\bar{H}}$ satisfy $|\lambda|L \ll 1$ where L is the length of the holographic interaction region then we find that

$$\vec{A} = (I + \vec{\bar{H}}L) \vec{A}_0 \quad (IV.9)$$

Thus a Fourier hologram written in the slab waveguide may perform a vector matrix multiplication taking each Fourier component of the incident field to each component of the output field. Under suitable conditions of mutual incoherence of the stored gratings and the fields vector matrix multiplication may also be shown to occur between the incident and output intensities.

In the architecture of figure IV.1 each Fourier component of the field incident on the the volume hologram corresponds to the field in a single channel of the input waveguide array. Similarly, each Fourier component of the field diffracted from the hologram corresponds to the field in a single channel of the output waveguide array. Since the hologram performs vector matrix multiplication between the incident and diffracted fields, the system as a whole performs vector matrix multiplication between the vector of field amplitudes in each input channel and the matrix of Fourier grating amplitudes in the volume hologram. Each component of the output vector is represented by a field amplitude in an output channel. Crosstalk, i.e. diffraction of one input channel by a grating corresponding to another input channel, and second order diffraction, are avoided by the Bragg wavevector matching condition given in equation (IV.5). As mentioned above, \vec{K}_g is the grating wavevector. \vec{k}_1 and \vec{k}_2 are constrained to lie on the normal surface. Since in an integrated geometry \vec{k}_1 and \vec{k}_2 are further constrained to the integrated plane, wavevectors satisfying the Bragg condition for a given \vec{K}_g are unique for an integrated hologram. Figure IV.2 shows graphically the relationship between the input vector, the output vector and the interconnection matrix. Each component of the input vector corresponds to a wavevector with an end point on the normal surface at the lower right. Each component of the output wavevector corresponds to a wavevector with an endpoint on the normal surface at the upper right. Each component of the interconnection matrix corresponds to a vector joining the endpoints of an input component and an output component on the normal surface. As long as each input and output wavevector differs by more than a vector of magnitude $\frac{1}{L}$ from the nearest neighboring wavevector, the input and the output vectors may be fully interconnected without crosstalk.

The requirement that the components of the input and output vectors be separated by $\frac{1}{L}$ arises from ambiguity in the grating wavevectors due to the finite volume of the hologram. This ambiguity limits the number, S , of independent interconnections which can be stored in a bulk volume hologram is limited to the number of resolvable spots in the volume, i.e.

$$S \approx \frac{V}{\lambda^3} \quad (IV.10)$$

where V is the volume of the recording medium and λ is the wavelength of the readout light. Psaltis *et.al.*¹¹ have shown that this limitation on the number of interconnections between two image planes limits the spots used for complete and independent interconnection of two planes to fractal grids of dimension $\frac{3}{2}$. In the case of the integrated volume holograms limitations of this sort on the input and output vectors due not arise. It can be shown that the number, S' , of interconnections which can be stored in an integrated hologram is

limited by ambiguity in Fourier space to

$$S' = \frac{A}{\lambda^2} \quad (IV.11)$$

where A is the area of the interaction region. There are of course further limits on S' arising from the finite bandwidth of the photorefractive response and the finite apertures of the optical system, but these limitations contribute only a scaling factor independent of A and λ . It is interesting to note that the integrated volume holographic vector matrix multiplier implements linear transformations with the same number of degrees of freedom as the bulk device using a two dimensional mask.

IV.3 Recording the Matrix

We now turn to the question of how the hologram representing \bar{H} may be recorded. Photorefractive effects offer a particularly simple and flexible approach. The photorefractive effect is a mechanism by which volume gratings may be created in waveguiding crystals such as LiNbO_3 , GaAs and $\text{SBN}^{2,3,4,5}$. Photorefractive gratings arise from the electrooptical modulation of the refractive index by the space charge field arising from the inhomogeneous charge distribution induced by the intensity distribution of the write beams. The amplitude of the photorefractive perturbation is proportional to the modulation depth, M , of the optical intensity. The diffraction efficiency of the resulting hologram is proportional to the square of the perturbation, i.e. to M^2 , and to the length of the interaction region. While the amplitude of the index perturbation which may be achieved in this way is small, the diffraction efficiency of photorefractive holograms is often large because of the volume nature of the perturbation region. The advantages of photorefractive gratings in the application described here are that they can be written with high resolution in real time with no substrate preparation or development steps, that they are erasable and modifiable, that writing photorefractive gratings does not add a bias to the mean index of refraction, and that the high sensitivity of the photorefractive effect to changes in wavelength allows the gratings to be written optically with short wavelengths and readout at longer wavelengths beyond the threshold for writing.

In order to have full independent control of the interconnection matrix it is necessary to separate the process of writing the hologram from the read-out process. This may be done by utilizing the rotational degeneracy of volume holograms. Since the right hand side of equation (IV.5) is unchanged by rotations about \bar{K}_g , an infinite set of coupled pairs \bar{k}_1 and \bar{k}_2 may be generated by rotating the normal surface about the grating wavevector. This degeneracy in the Bragg condition allows us to write gratings in integrated waveguides with unguided light (radiation modes) which may be readout by guided beams. If we include in our analysis a change of the magnitudes of the optical wavevectors due the transition from guided to unguided beams and a difference in the wavelengths of the write and read beams, then it may be shown that guided beams at wavelengths beyond the threshold of the photorefractive response may be diffracted by holograms written by

unguided light at a shorter wavelength. This arrangement allows us to take advantage of the long interaction lengths, high beam intensities, and compact construction of integrated devices while avoiding optical damage and photorefractive scattering due to the readout beams. Since we have all the degrees of freedom of bulk optics in positioning the write beams, we will be able to describe below a simple method for writing with a single exposure a hologram which completely and independently interconnects every resolvable spot at the input to the waveguide with every resolvable spot at the output.

Vector matrix multiplication is performed in the architecture we are proposing by using $n_1 n_2$ gratings to map n_1 input frequencies to n_2 output frequencies. Each grating may be associated with a pair of radiation modes by rotating the input and output beams about the grating vector. Figure IV.3(a) is a wave matching diagram which graphically shows the Bragg matching condition as a requirement that the end points on the normal surface of coupled optical wavevectors be joined by the grating wavevector. Figure IV.3(b) shows a pair of unguided optical beams which may be associated with this wavevector. Figure IV.3(c) shows a second pair of optical beams generated upon rotation about a second wavevector coupling one of the original guided beams with a third guided beam. As can be seen in figure IV.3(d) the unguided wavevectors into which the guided beam is rotated are different for the two different grating wavevectors. This simple graphical demonstration points out a problem which can be shown to hold in general in the architecture we are proposing, which is that as many as $2n_1 n_2$ distinct radiation modes must be available to write the $n_1 n_2$ gratings coupling $n_1 + n_2$ guided modes. One means of presenting all these radiation modes simultaneously in each write cycle might be to use a second volume hologram to store the $2n_1 n_2$ beams.

The amplitude of a photorefractive index perturbation is proportional to the modulation depth of the writing beams. For the case of writing with $n_1 n_2$ beams of approximately equal intensities, the modulation depth of each grating will be proportional to $(n_1 n_2)^{-1}$. The modulation depths of the individual gratings are also proportional to $(n_1 n_2)^{-1}$ for sequential writing if the asymmetry between the erase and write times is not large. By using the architecture of figure IV.4 it is possible to reduce this factor to $(\sqrt{n_1 n_2})^{-1}$ by writing all the gratings with a single reference beam. In this geometry the modulation depth of the $(ij)^{th}$ grating is

$$M_{ij} = \frac{\sqrt{I_{ij} I_R}}{I_R + N \langle I_{ij} \rangle} \quad (IV.12)$$

Where I_{ij} and I_R are the intensities of the $(ij)^{th}$ write beam and of the reference respectively. $\langle I_{ij} \rangle$ is the mean value of the write beam intensities. The expected value of M_{ij} has a maxima of $(2\sqrt{n_1 n_2})^{-1}$ when $\langle I_{ij} \rangle = \frac{I_R}{\sqrt{n_1 n_2}}$.

A problem with the architecture shown in figure IV.4 arises from the fact that it is not possible to write an interconnection matrix in the plane using a single reference while still satisfying the Bragg condition for each interconnection. This problem may be overcome by taking advantage of the fact that the interaction region of the hologram is very thin out of the plane of the waveguide. If the modes coupled by a given grating are well guided then the coupling efficiency is insensitive to Bragg mismatch out of the waveguiding plane

if the phase of the grating remains constant within the coupling region. This fact allows us to have a Bragg mismatch out of the plane of the waveguide. This concept is illustrated in figure IV.5. In figure IV.5(a) we show a set of fringes for a grating written with \vec{K}_g confined to the surface of the waveguide. Note that the phase of the grating is constant along the vertical direction. In figure IV.5(b) \vec{K}_g is not confined to the waveguiding plane and the phase of the grating is not constant along the vertical direction. However, since the interaction region for well confined modes includes only the waveguide itself, high diffraction efficiencies may be achieved between guided modes if the phase of the grating is constant in the vertical direction in a region approximately limited to the waveguide itself. This condition may be expressed in the form

$$(\vec{K}_g)_z d \ll 1 \quad (IV.13)$$

where d is the thickness of the interaction region. Since d may be very small for an optical waveguide, we find that our architecture will tolerate a fairly large Bragg mismatch out of the waveguiding plane. This allows us to write the full interconnection matrix with a single reference as shown in figure IV.4.

It is still necessary to fabricate the mask T in such a way that Bragg matching in the plane of the waveguide is achieved for each interconnection. The Fourier lens L_1 is aligned so that a pixel on its optical axis is collimated such that

$$\kappa_{00x} = K_x; \kappa_{00y} = K_y = 0; \kappa_{00z} = -K_z \quad (IV.14)$$

where $\vec{\kappa}_{00}$ and \vec{K} are the wavevectors of the collimated beam and of the reference respectively. The requirement that the grating produced by the reference and the light from pixel ij be the grating that connects the i^{th} input channel and the j^{th} output may be expressed $\vec{\kappa}_{ij} = \vec{K} - \vec{k}_{ij}$. Since \vec{k}_{ij} has no component out of the plane of the waveguide, the \hat{e}_z component of this equation is satisfied if, as we saw above,

$$(\vec{K} - \vec{\kappa}_{ij})_z \ll \frac{1}{d} \quad (IV.15)$$

In the paraxial approximation this requirement becomes $\frac{\rho}{f} \ll \sqrt{\frac{\lambda}{d}}$ where ρ is the greatest distance from the optical axis of a pixel on T . Having satisfied this constraint, the \hat{e}_y and \hat{e}_z components of the phase matching equation may be satisfied by properly placing the pixels on the mask. An example of a suitable mask for a four by four device with a 400 μm separation between channels is shown in figure IV.6. We have assumed a lens of focal length 25cm between the mask and the waveguide and integrated lenses of focal length 1cm. The angle between the writing beams is assumed to be three degrees.

IV.4 Experimental Results and Discussion

We have written photorefractive gratings in single mode titanium indiffused slab waveguides on nominally pure y cut LiNbO₃. The gratings were written using unguided

light at 488 nm from an Ar^+ laser. The red HeNe line was used for readout. The grating wave vectors were nearly parallel to the c axis to make use of the r_{33} electrooptic coefficient and the photovoltaic field. Optical propagation was along \hat{e}_y for the write beams and along \hat{e}_z for the guided read beams. The effective index for the guided HeNe mode was 2.248, while the extraordinary index of the substrate at 488 nm is about 2.6. Thus, between the write and read beams there was an expansion of the angle between the beams Bragg matched to a particular grating by a factor of about 1.5. Small angles were used because our waveguide was not guiding along \hat{e}_z . We have obtained coupling efficiencies of up to 4.5% between guided modes of interaction lengths of about 1 cm. We have also been successful at coupling a single guided mode into two diffracted modes, though we have not yet implemented the mask scheme described above to achieve multibeam coupling with high efficiency. Photographs of zero and first order spots diffracted from the end of the waveguide in the one and two grating cases are shown in figure IV.7. In two grating case both gratings were written with a single reference beam. Both are slightly Bragg mismatched for readout.

Various problems arise in the implementation of the architecture we have proposed. Two specific problems concern the depletion of the pump beams and the implementation of the matrix in grating amplitudes. Since the interaction region for integrated volume holograms may be very long, the coupling efficiency of these holograms may be quite high. In the depleted pumps regime the accuracy to which the device represents a true vector matrix multiplier will be compromised. The extent to which this is a problem depends upon the application envisioned. Since the amplitude of each grating is linear in modulation depth, which is a nonlinear function of the total background intensity, it is only possible in a statistical sense to associate a given grating amplitude with a specific transmittance in the writing mask. For small scale applications both of these problems may be surmounted by considering the expected signals from each channel. For larger scale applications, the application must be tolerant to these nonlinearities. Optical neural computers⁸ are one such application.

A third problem arises from the effect of titanium on the photorefractive properties of LiNbO_3 . Glass *et. al.* found that Ti doping increases the dark conductivity of LiNbO_3 , thereby reducing the grating storage time⁵. In our waveguides we found grating lifetimes of several days under continuous guided readout with about 50 μW . One method of increasing this lifetime might be to use proton exchanged waveguides⁹. In applications where storage time is not a concern, GaAs or SBN might be used as a substrate.

IV.5 conclusion

In conclusion, we have proposed an architecture using coupling from a set of input channel waveguides to a set of output channels using multiwave photorefractive mixing in a slab waveguide to implement vector matrix multiplication. We have demonstrated the feasibility of out of plane photorefractive deflection of guided beams. By using out of plane beams to address the nonlinearity we are able to combine the holographic capabilities of

photorefractive crystals with the capabilities of integrated optics.

References

- [1] C.S. Tsai, *IEEE Trans. Circuits and Systems* CAS-26,1074(1979)
- [2] R.P. Kenan, D.W. Vahey, N.F. Hartman, V.E. Wood, C.M. Verber, *Opt. Eng.* 15,12(1976)
- [3] O.V. Kandidova, V.V. Lemanov, B.V. Sukharev, *Sov. Phys. Tech. Phys.* 29,1019 (1984)
- [4] W.S. Goruk, P.J. Vella, R. Normandin, G.I. Stegeman, *Appl. Opt.* 20,4024(1981)
- [5] A.M. Glass, I.P. Kaminow, A.A. Ballman, D.H. Olson, *Appl. Opt.* 19,276(1980)
- [6] N.V. Kukhtarev, V.B. Markov, S.G. Odulov, M.S. Soskin V.L. Vinetskii, *Ferroelectrics* 22,949(1979)
- [7] D.Y. Zang, C.S. Tsai, *Appl. Opt.* 25,2264,(1986)
- [8] D. Psaltis and N. Farhat, *Opt. Lett.* 10,98(1985)
- [9] J.L. Jackel, C.E. Rice, J.J. Veselka, *Appl. Phys. Lett.* 41,607(1982)
- [10] J. Goodman, A.R. Dias, and I.M. Woody, *Opt. Lett.* 2,1(1978)
- [11] D. Psaltis, J. Yu, X. G. Gu, and H. Lee, Second Topical Meeting on Optical Computing, Incline Village, Nevada, March 16-18,1987

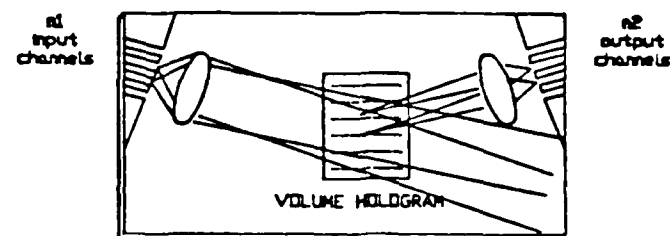


Fig. IV.1. Integrated optical matrix vector architecture.

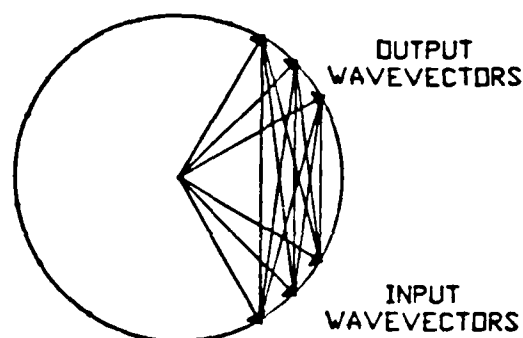


Fig. IV.2. Wavematching between the input, the output, and the interconnection matrix.

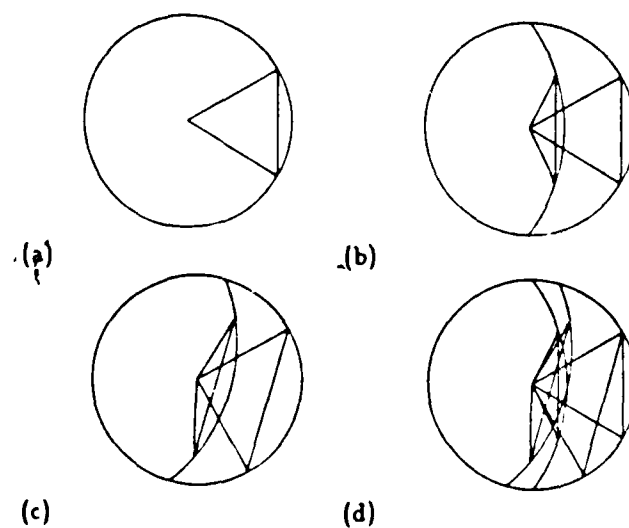


Fig. IV.3. Rotational degeneracies of Bragg matched holograms.

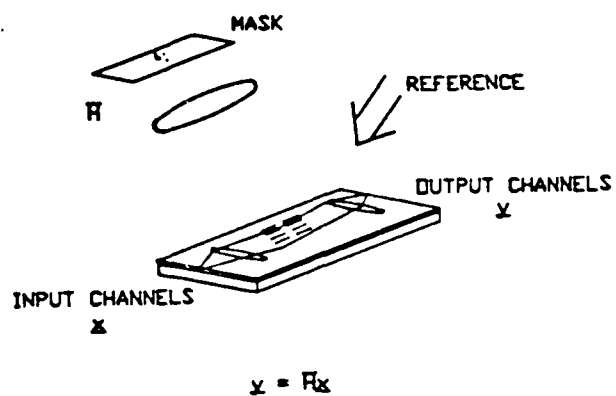


Fig. IV.4. Recording with unguided light and a single reference.

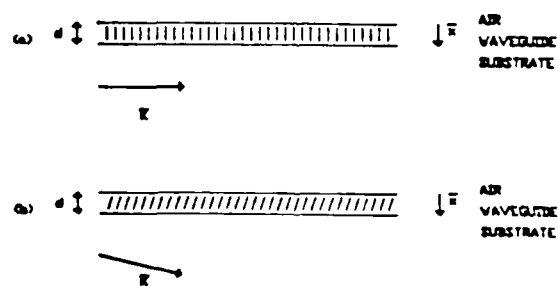


Fig. IV.5. Bragg mismatch out of the guiding plane.

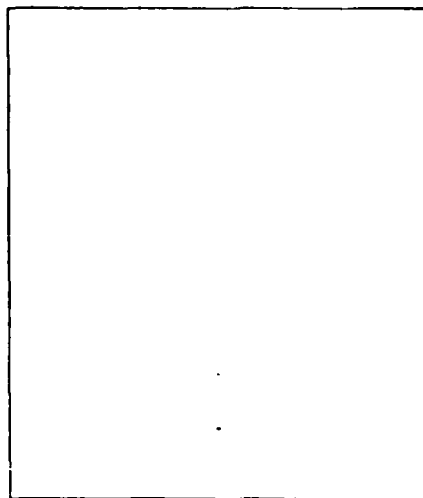


Fig. IV.6. Mask for recording a 4 x 4 matrix.



Fig. IV.7. Diffracted and undiffracted beams in LiNbO_3
 One grating. Two gratings.

END

DATE

FILMED

8-88

DTIC